

Your grade: 100%

Your latest: 100% • Your highest: 100% • To pass you need at least 80%. We keep your highest score.

Next item →

1 / 1 point

1. The Lunar Lander is a continuous state Markov Decision Process (MDP) because:
- ☐ The state-action value  $Q(s, a)$  function outputs continuous valued numbers
  - ☒ The state contains numbers such as position and velocity that are continuous valued.
  - ☐ The reward contains numbers that are continuous valued
  - ☐ The state has multiple numbers rather than only a single number (such as position in the  $x$ -direction)

✔ Correct  
That's right!

1 / 1 point

2. In the learning algorithm described in the videos, we repeatedly create an artificial training set to which we apply supervised learning where the input  $x = (s, a)$  and the target, constructed using Bellman's equations, is  $y =$  \_\_\_\_\_?
- ☒  $y = R(s) + \gamma \max_{a'} Q(s', a')$  where  $s'$  is the state you get to after taking action  $a$  in state  $s$
  - ☐  $y = \max_{a'} Q(s', a')$  where  $s'$  is the state you get to after taking action  $a$  in state  $s$
  - ☐  $y = R(s)$
  - ☐  $y = R(s')$  where  $s'$  is the state you get to after taking action  $a$  in state  $s$

✔ Correct

1 / 1 point

3. You have reached the final practice quiz of this class! What does that mean? (Please check all the answers, because all of them are correct!)

✔ The DeepLearning.AI and Stanford Online teams would like to give you a round of applause!

✔ Correct

✔ What an accomplishment -- you made it!

✔ Correct

✔ Andrew sends his heartfelt congratulations to you!

✔ Correct

✔ You deserve to celebrate!

✔ Correct